

# Computational Research Division Report

## Supercomputing Report Card

Assessment of applications performance on six systems identifies petascale challenges

A comprehensive supercomputer performance evaluation undertaken by CRD scientists has won a Best Paper award in the Application track at the IEEE International Parallel and Distributed Processing Symposium (IPDPS), underscoring the significant contribution the research makes to improving scientific applications for the arrival of petascale computing.

The paper, "Scientific Application Performance on Candidate PetaScale Platforms," is an accumulation of four years of exhaustive study that examined the performance of six codes, representing a wide range of research disciplines, on six supercomputers throughout the country.

Headed by Lenny Oliker, the research not only answered key questions about high-performance systems, it also identified the shortcomings to overcome in order to scale the codes to run on next-generation supercomputers.

"If you go to the literature today and

look for comparisons on how the machines differ from one another on realistic, large-scale applications, there is surprisingly little information out there," Oliker said. "The research benefits everyone from application scientists to next-generation supercomputer designers."

The research is a collaboration between computer and application scientists at CRD and NERSC. The co-authors are Andrew Canning, Jonathan Carter, Costin Iancu, Michael Lijewski, Shoaib Kamil, John Shalf, Hongzhang Shan and Erich Strohmaier. Stephane Ethier from the Princeton Plasma Physics Laboratory and Tom Goodale from Louisiana State University also contributed to the work.

The IEEE symposium is scheduled to honor the paper's authors and have them present the paper on March 28 in Long Beach, California.

"This is indeed an outstanding accomplishment," said Horst Simon, Associate  
*continued on page 2*

## BETTER ENVIRONMENT

New server enables scientists to access and analyze water data in minutes instead of days

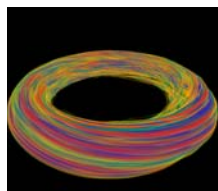
A collaboration among Microsoft, Berkeley Lab and UC Berkeley is underway to develop a scientific data server for amassing and organizing water data from diverse sources, a system that will accelerate research in the increasingly important areas of water and climate change.

Called Microsoft e-Science, the project is part of the Berkeley Water Center's effort to marshal expertise from public institutions and the private sector and support projects that enable researchers to easily access and work with water data. The one-year-old center is the brainchild of Berkeley Lab and UC Berkeley's College of Engineering and College of Natural Resources.

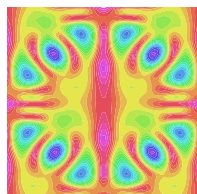
Local, state and federal governments have long collected detailed information about water supplies, such as measuring the river flows and water content in the Sierra's winter snows. They use the data

*continued on page 3*

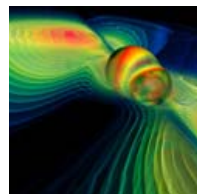
## Visualization of Evaluated Simulations



GTC electrostatic potential field



ELBM3D vorticity turbulence



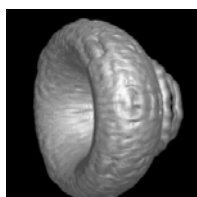
Cactus black hole collisions



Beambeam3D particle tracks



PARATEC CdSe quantum  
dot electron state



HyperCLaw helium bubble  
deformation

## Microbial Talent

Enhanced metagenome data system will aid research on biofuels, wastewater treatment

The DOE Joint Genome Institute (JGI), collaborating with CRD's Biological Data Management and Technology Center (BDMTC), announced last month an upgrade of the metagenome data management and analysis system, IMG/M, on the anniversary of its launch.

Developed by BDMTC's data management and software engineers and scientists of JGI's Genome Biology and Microbial Ecology Programs, IMG/M is a valuable asset for a growing community

*continued on page 4*

## SIAM Postscript

This year's SIAM Conference on Computational Science and Engineering (CSE) featured a strong showing from CRD scientists, who spent a week in Costa Mesa, California this month to discuss topics that ranged from nanoscience simulations to scientific data management.

In all, 13 researchers from the division attended the week-long conference, which also ran concurrently with the SIAM Workshop on Combinatorial Scientific Computing.

"The large number of contributions to one of the key conferences in the field shows the breadth of research and expertise we have in computational science and engineering," said **Horst Simon**, Associate Lab Director of Computing Sciences and one of the conference attendees.

SIAM's Activity Group on Computational Science and Engineering—currently chaired by CRD's **John Bell**—organized this biennial conference. Bell also is chairing a session on fluid dynamics at the conference.

CRD scientists organized mini-symposiums, presented papers and discussed posters during the conference and workshop:

**Arie Shoshani** was invited to speak on "Scientific Data Management: Essential Technology for Data-Intensive Science." Shoshani talked about technological advancements that allow researchers to effectively mine for and analyze data when they have to wade through very large datasets. He also organized the mini-symposium "Data Management for Scientific Applications," which described the technology for, among other things, high bandwidth parallel file systems, efficient parallel statistical computing and automation of large scientific workflows.

**Juan Meza** discussed "Surface Structure Determination of Nanostructures Using a Mesh Adaptive Optimization Method." The work examined a method of determining the atomic configuration of a nanostructure's surface that serves as a cost-effective alternative to the common low-energy electron diffraction (LEED) method. Meza proposed the use of generalized pattern search methods in combination with a simplified physics surrogate for the full fidelity physics model.

**John Shalf** from NERSC and **Lenny**

## Report card *continued from page 1*

Lab Director of Computing Sciences and head of CRD and NERSC. "Getting into the program is already a big deal. Getting best paper is exceptional."

The project traces its genesis to the 2002 birth of Earth Simulator, which reigned on the TOP500 list for two and half years and upstaged the supercomputing community in the United States. Olikier and his team were among the small number of U.S. scientists given access to Earth Simulator and traveled to Japan several times to evaluate the system. The researchers, including Ethier, have reported their findings in over 10 technical papers.

The Earth Simulator evaluation prompted Olikier and other American researchers to kick-start the project to compare performances of widely used supercomputers in national labs across the United States. The team set out to determine how well the codes currently used by scientists would fare on the supercomputers and the tradeoffs among various system designs.

Olikier's team chose six codes that represented a broad spectrum of research areas: magnetic fusion (GTC), fluid dynamics (ELBM3D), astrophysics (Cactus), high energy physics (BeamBeam3D), materials science (PARATEC) and AMR gas dynamics (HyperCLaw).

They ran these six codes on each of the six supercomputers that, in turn, represented a wide range of architectures. The systems were Bassi and Jacquard from Lawrence Berkeley National Laboratory, Jaguar and Phoenix from Oak Ridge National Laboratory, Blue Gene/L from Argonne National Laboratory and another Blue Gene/L from IBM Thomas J. Watson Research Center.

Bassi is an IBM Power5 system with 888 compute processors (111 8-way nodes) that runs on AIX. Jacquard contains 640 single-core AMD Opteron processors (320 2-way nodes), running Linux 2.6.5, while Jaguar features 14,400 dual-core Opteron processors (5,200 2-way nodes) running Catamount. Pheonix is a vector-based Cray X1E system contain-

ing 768 processors (96 8-way MSP nodes) and runs UNICOS.mp. The Blue Gene/L at Argonne is an IBM PowerPC 440-based system with 2,048 processors (1024 2-way nodes) and runs SuSE Linux OS (SLES9). The Blue Gene/L at IBM's research center contains 40,000 processors.

The study produced results that showed the strengths and weaknesses of various



**Lenny Olikier**

high-performance systems. For example, the Power5-based Bassi achieved the highest per processor raw performance running four of the six codes. The X1E-based Pheonix system, on the other hand, produced impressive raw per-

formance on GTC and ELBM3D. However, applications with nonvectorizable portions didn't perform as well on Phoenix as a result of the "imbalance between the scalar and vector processors," the researchers said.

"Our results indicate that our evaluated codes have the potential to effectively utilize petascale resources," Olikier and others wrote in the paper. "However, several applications will require re-engineering to incorporate the additional levels of parallelism necessary to achieve the vast concurrency of upcoming ultra-scale systems."

The team identified PARATEC and BeamBeam3D as among the applications that would benefit from retooling.

The scientists' findings have been presented at the SC03, SC04, and SC05 conferences, as well as in the recently published SIAM book "Parallel Processing for Scientific Computing."

Previous research on scientific application behavior for differing supercomputers also garnered Olikier and other researchers best-paper awards in SC99 and SC2000.

You can read more about the paper that will be presented next month at <http://crd.lbl.gov/~oliker/papers/ipdps07.pdf>.

**Olikier** organized the mini-symposium "Beyond Petaflops: Specialized Architectures for Power Efficient Scientific Computing." The discussion examined the potential of custom supercomputing

platforms using embedded systems technology, which holds the promise of multi-petaflop computing at a fraction of the cost and power consumption of the con-

*continued on page 5*



## Water data server *continued from page 1*

to make allocation decisions for farms, businesses and residential consumers. However, these agencies use different methods to collect and archive the information, posing a challenge for scientists who need to retrieve and integrate all those data sets in order to carry out comprehensive analyses.

The e-Science project strives to ease that headache for scientists. The project team already has developed a prototype data server, which runs on Microsoft SQL Server 2005. The team is now testing the system by loading data about Northern California's Russian River watershed from a myriad of agencies.

"Because of the differences in the data, the loading of each data file presents a new challenge, and matching data across different data sets is difficult," said Deb Agarwal, head of CRD's Distributed Systems Department and the Berkeley Water Center's IT Advisor. "There is a perception that once the data is in an archive, science is enabled on a grand scale. But data availability is only the first step in the process."

Agarwal and other project researchers already have demonstrated the prototype server to the scientific community. For example, at the FLUXNET Synthesis Workshop in Italy this month, project team members Matt Rodriguez, a CRD scientist, and Catharine van Ingen from Microsoft ran the data server to show how scientists could find and plot cross-network data in minutes, rather than days. The data to be analyzed were 600

site-years of data, most of which had not been used before in cross-site analysis. Through use of the data server, scientists can spend time exploring the data rather than collating them.

At the European Geosciences Union General Assembly 2007 in April, Agarwal, van Ingen and Dennis Baldocchi from UC Berkeley will discuss the server and their support of its users in a paper titled "A Next Generation Flux Network Data Server."

Microsoft's support is critical for the project because approximately 90 percent of the researchers accessing these data archives are working on Windows-based desktop computers. Van Ingen brings expertise from her work as an engineering professor and software

---

*"We are developing tools that we hope will make this data server a natural research tool, a kind of 'collaborative data server in a box' for science."*

---

— Deb Agarwal

---

expert, as well as a Microsoft insider who knows where to turn for help in the company.

Developing the prototype server was an important milestone for the project. To build it, the project team started with the data archive of the AmeriFlux network of 149 research towers located around the Americas.

Using arrays of sensors, the towers provide continuous observations of ecosystem-level exchanges of CO<sub>2</sub>, water and energy, essentially recording how the ecosystem "breathes." The AmeriFlux archive currently contains 192 million data points stored as hundreds of flat files.

Researchers analyzing this data currently download a copy of the data for use in local analysis. Since the data is continually being updated and corrected,

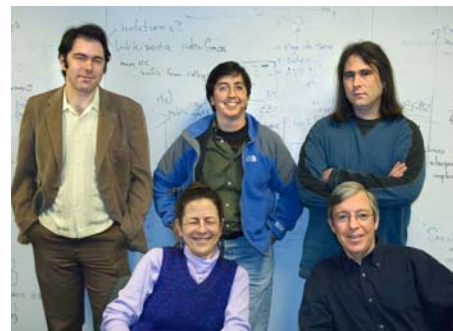


Photo by Gordon Bell of Microsoft

The e-Science team is developing better tools for environmental research. In the front row (from left to right) are Catharine van Ingen and Stuart Ozer, both of Microsoft. Back row: Matt Rodriguez, Deb Agarwal and Monte Goode, all from CRD.

each researcher typically ends up with a different version of the data.

Agarwal and her staff, Rodriguez and Monte Goode, working with van Ingen and another expert from Microsoft, Stuart Ozer, designed the server to make the AmeriFlux data easier to use. The approach incorporated a database and a "data cube," a type of database structure optimized for data mining.

"Although the long-term plan is to develop the server for use in gathering and understanding watershed data, using the AmeriFlux archive data gave us an excellent environment to use as we designed the server," Agarwal said.

At the annual AmeriFlux meeting last October, Gretchen Miller, a UC Berkeley graduate student working with the BWC, demonstrated an AmeriFlux data cube, and Agarwal reported on the data server project. The presentation was followed by a discussion of the project's next steps and a request for the e-Science team to support the FLUXNET meeting.

While developing the server is a major part of the project, the long-term goal is to develop a portable system that can be maintained by the researchers themselves.

"Right now we're at the edge of computer science and research, where we are developing tools that we hope will make this data server a natural research tool, a kind of 'collaborative data server in a box' for science," Agarwal said.

You can learn more about the Berkeley Water Center at <http://www-esd.lbl.gov/BWC/>.



The Russian River watershed provides water for a booming wine grape region that also is growing in human population. The e-Science team in adding the watershed data to its prototype server.

## Metagenome *continued from page 1*

of biologists. Metagenomics refers to the genomic study of batches of microbes in their natural environments instead of those cultivated individually in lab cultures. IMG/M provides tools for analyzing the capability of microbial communities based on their metagenome sequence—in the context of reference isolate genomes—using a variety of public functional and pathway resources.

Since the release of its initial experimental version in 2006, IMG/M has been used by JGI scientists and their collaborators for analyzing several environmental microbial communities, and has been demonstrated at professional meetings and conferences. The new version of IMG/M will be presented at JGI's second annual user meeting on March 28 in Walnut Creek, California.

The enhanced version of IMG/M now offers aggregate genome (metagenome) data generated from microbial community samples that have been the subject of recently published studies. These include samples from biological phosphorus removing sludge (*Nature Biotechnology* Volume 24, Number 10, October 2006), human distal gut (*Science* 312: 1355-1359, 2 June 2006), a gutless marine worm (*Nature* 443, 950-955, 26 October 2006) and obese and lean mouse gut (*Nature* 444, 1027-131, 21 December 2006).

These research projects give scientists valuable clues to how these microbes could help clean up the environment and produce alternative transportation fuels. For example, the phosphorus-removing sludge study, the genome sequence of bacteria *Accumulibacter phosphatis* is enabling researchers to figure out how the bugs could accumulate massive amounts of phosphorus inside their cells. Being able to cultivate the microbes, which currently are impossible to grow in laboratory culture, can play a key role in reducing phosphorous during wastewater treatment and protecting rivers and lakes from excessive algae blooms.

In addition, IMG/M includes three of the simulated metagenome data sets employed for benchmarking several assembly, gene prediction and binning methods (see <http://fames.jgi-psf.org/>).

"This is the first production version of IMG/M," said Victor Markowitz, head of BDMTC. IMG/M's reference isolate

## Hall of Fame



Vern Paxson

Vern Paxson, a researcher in CRD's Distributed Systems Department, has been named a fellow of the Association for Computing Machinery (ACM). The fellowship is given to scientists who have made "outstanding technical and professional achievements in the field of information technology," according to ACM.



Horst Simon

Paxson is probably best known for his original development work on Bro, the Lab's intrusion detection system, which monitors incoming and outgoing traffic and alerts cyber security staff when suspicious traffic patterns are detected. Paxson will be inducted at the ACM Awards Banquet on June 9 in San Diego.

Energy Secretary Samuel Bodman has appointed Horst Simon, head of Computing Sciences, to the Advanced Scientific Computing Advisory Committee (ASCAC), which is the federal advisory committee to the Office of Advanced Scientific Computing Research. More information about ASCAC can be found at <http://www.sc.doe.gov/ascr/ASCAC/ASCAC.html>.



Juan Meza

Juan Meza, head of CRD's High Performance Computing Research Department, has been named to two committees of two research societies. Cleve Moler, who begins his term as president of the Society for Industrial and Applied Mathematics this month, asked Meza to serve a three-year term on the SIAM Committee on Science Policy starting Jan 1, 2007. More information about the committee can be found at <http://www.siam.org/about/science/sci-pol.php>.



Philip Sternberg

Meza was also elected to a three-year term on the American Association for the Advancement of Science (AAAS) Electorate Nominating Committee of the Section on Mathematics. The 24 sections of AAAS arrange symposia for the annual meeting, elect officers and provide expertise for association-wide projects.

Philip Sternberg has joined CRD's Scientific Computing Group (SCG) as a postdoctoral fellow. He will be working with Chao Yang and Esmond Ng on a single particle reconstruction project and a nuclear physics SciDAC project. Sternberg received his Ph.D. in mathematics from UC Davis last year and a bachelor's degree in mathematics from UC Berkeley in 2002. He was a student intern in the SCG last summer.

genomes were included from version 2.0 of JGI's Integrated Microbial Genomes (IMG) system, a total of 2,301 isolate genomes consisting of 595 bacterial, 32 archaeal, 13 eukaryotic, and 1,661 virus genomes. The Computational Biology and Bioinformatics group at Pacific Northwest National Laboratory (PNNL) carried out the large-scale gene similarity computations on the Molecular Sciences Computing Facility supercomputer.

"The support provided by our colleagues at PNNL has been critical in building IMG/M—computations that otherwise would have taken several weeks were completed in about a day," said Markowitz.

BDMTC has expertise in the development of biological data management systems and bioinformatics tools. BDMTC is collaborating with DOE JGI on the development and maintenance of IMG

*continued on page 5*



## SIAM *continued from page 2*

ventional HPC design.

As part of this track, Olikar presented a talk titled “Towards Ultra-High Resolution Models of Climate and Weather.” He discussed the use of embedded processor technology for building a supercomputer that is specifically designed to create kilometer-scale atmospheric models. Generating those simulations would require a sustained 10-petaflop computer, and using customized processors can lead to a more power-efficient system.

**Ali Pinar** delivered a talk titled “Vulnerability Analysis on the Electric Power Grid.” Pinar talked about a mixed integer nonlinear formulation for the vulnerability analysis of electric power systems, and how the special structure of this problem can be exploited to further reduce complexity.

**Andrew Canning** discussed “New Eigensolvers for Large Scale Nanoscience Simulations.” He presented a comparison of some different eigensolvers for the solution of electronic structure calculations using semi-empirical potentials. He illustrated the approach with some applications in nanoscience, such as quantum dots and quantum wires.

**Phil Collela** moderated the panel “Graduate Education Session: Career Path in Computational Science and Engineering – Part II.”

**Xiaoye “Sherry” Li** presented the work “Extra-Precise Iterative Refinement for Least Squares Problems.” Linear least squares (LLS) fitting is the most widely used data modeling technique and is included in almost every data analysis

system (e.g., spreadsheets). With limited use of extra precision, for all but the most ill-conditioned LLS problems, the technique can eliminate the concerns of the floating-point calculation errors present in the solution. Li also chaired a track titled “Spare Direct Methods.”

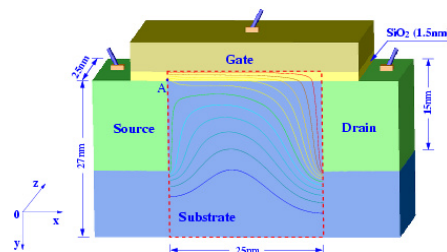
**Panagiotis Stinis** chaired the track on fluid dynamics and gave a talk on “Long Memory Mori-Zwanzig Models for the Euler Equations.” The talk presented a new class of reduced models for the 3D Euler equations. The models are based on the long memory caused by the appearance of organized vortical structures.

**Christof Voemel** presented “The Use of Bulk Information to Improve the Scalability of Parallel Band Gap Computations for Quantum Dots.” The calculation of photoluminescence properties of semiconductor quantum dots can be accelerated three- to four-fold by exploiting the physical relationships between the interior of a dot and an ideal crystal. This talk described the corresponding computational procedure in the framework of preconditioned parallel iterative eigensolvers for interior eigenstates of the quantum dot Hamiltonian.

**Esmond Ng**, a featured speaker in the Combinatorial Scientific Computing workshop, gave the talk, “Towards Optimal Petascale Simulations.” The SciDAC-funded project aims to develop scalable solvers that will allow large-scale scientific simulation codes to run on petascale computer architectures.

Two CRD scientists, **Lin-Wang Wang** and **Zhengji Zhao**, created posters for the conference. Canning presented Wang’s poster titled “Atomistic Pseudopotential Simulation of Nanometer Sized CMOS Devices.” Meza presented Zhao’s poster titled “Linear Scaling 3D Fragment Method for Petascale Nanoscience Simulations.” The poster described an

## Lin-Wang Wang’s Poster



**Linear combination of bulk band (LCBB)** is used to simulate the quantum mechanical effects of a CMOS device consisting of millions of atoms. Empirical pseudopotentials are used to describe the single particle Hamiltonian. The single electron wavefunctions are solved using the LCBB method. Potential selfconsistency is achieved by solving the Poisson equation with the occupied charge density. Approximation formulas are proposed to calculate the occupied charge density and electron current under a nonequilibrium condition.

O(N) method that is well suited for petascale nanoscience simulations. The method divides a large system into small fragments. It solves the electronic wavefunctions of each fragment independently, and then it patches the results together.

## Metagenome *continued from page 4*

(<http://img.jgi.doe.gov>) and IMG/M (<http://img.jgi.doe.gov/m>).

JGI, supported by the DOE Office of Science, unites the expertise of five national laboratories—Lawrence Berkeley, Lawrence Livermore, Los Alamos, Oak Ridge, and Pacific Northwest—along with the Stanford Human Genome Center to advance genomics in support of the DOE mission related to clean energy generation and environmental characterization and clean-up.

More information about BDMTC can be found at <http://crd.lbl.gov/html/BDMTC/index.html>.

## About CRD Report

CRD Report, which publishes every other month, highlights the cutting-edge research conducted by staff scientists in areas including turbulent combustion, nano materials, climate change, distributed computing, high-speed networks, astrophysics, biological data management and visualization. Find previous CRD Report articles at <http://crd.lbl.gov/html/news/CRDreport.html>.

### DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California. Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.